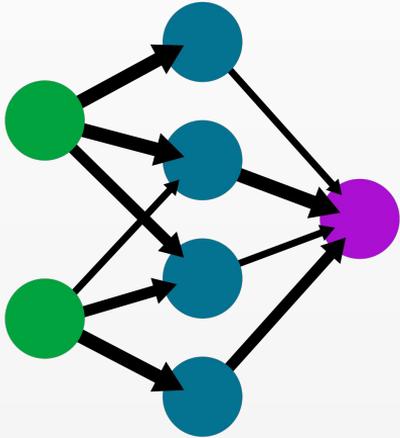


# MODELADO DE LA MASA NUCLEAR A TRAVÉS DE REDES NEURONALES

---

Sofía Anna Scozziero



# OUTLINE

01 Modelos Nucleares

02 Generalidades de RN

03 Complementos de RN

04 Paper de Mumpower

01

---

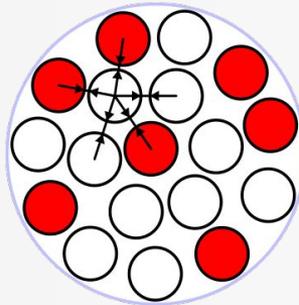
# MODELOS NUCLEARES

# PROBLEMA

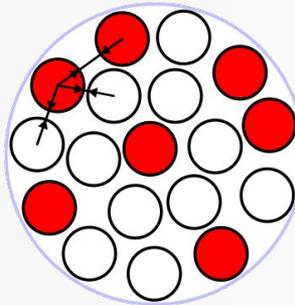
¿Podemos predecir la masa (entre otras propiedades) de un núcleo a partir de su número de neutrones y protones?

# MODELO DE LA GOTA LÍQUIDA

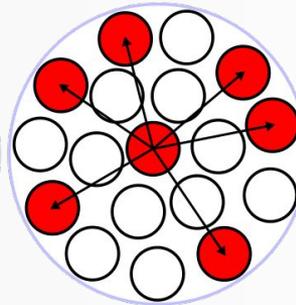
---



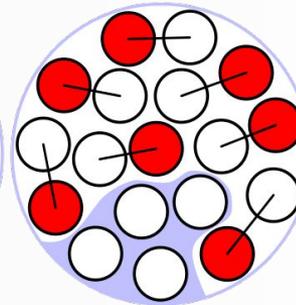
Volume



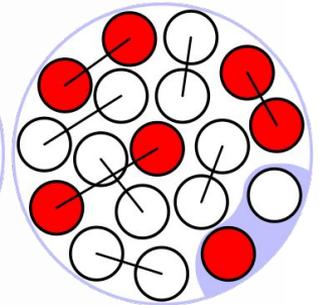
Surface



Coulomb



Asymmetry



Pairing

# MODELO DE LA GOTA LÍQUIDA

$$m(Z, N) = Zm_p + Nm_n - \frac{BE(Z, N)}{c^2}$$

$$BE(Z, N) = \alpha_v A - \alpha_s A^{2/3} - \alpha_C \frac{Z^2}{A^{1/3}} + \alpha_a \frac{(N - Z)^2}{A} + \alpha_p \delta(Z, N)$$

$$\delta(Z, N) = \begin{cases} 1 & Z, N \text{ pares} \\ 0 & Z \text{ par } N \text{ impar o } Z \text{ impar } N \text{ par} \\ -1 & Z, N \text{ impares} \end{cases}$$

# MODELO DE LA GOTA LÍQUIDA

---

A favor

- Permite calcular las energías de ligadura y explica las reacciones nucleares

En contra

- Es semi empírico
- No ajusta bien a todos los núcleos
- No explica otras propiedades (ej: los números mágicos o los estados excitados)

# MODELO DE CAPAS

Hamiltoniana many  
body

$$\hat{\mathcal{H}}\psi = E\psi$$

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_A)$$

Aproximación de campo medio  
(nucleones independientes)

$$\left[ \frac{-\hbar^2}{2m} \nabla^2 + U \right] \phi_\alpha = \epsilon_\alpha \phi_\alpha$$

potencial medio

$$U(\mathbf{r})$$

funciones de onda de  
una partícula indep

$$\phi_\alpha(\mathbf{r})$$

# MODELO DE CAPAS

---

Separación de variables en parte radial y armónicos esféricos (potencial central)

$$\psi(\mathbf{r}) = \frac{u(r)}{r} Y(\phi, \theta)$$

$$\left[ \frac{-\hbar^2}{2m} \frac{d^2}{dr^2} + \frac{\hbar^2}{2m} \frac{\ell(\ell+1)}{r^2} + U(r) \right] u(r) = Eu(r)$$

# MODELO DE CAPAS

En una primera aproximación se puede usar el oscilador armónico isotrópico

$$U(r) = \frac{1}{2}m\omega^2 r^2$$



En un modelo de capas (u orbitales) cerradas, las configuraciones más estables son aquellas que completen los niveles de energía.

Entonces los valores de N o Z más estables (números “mágicos”) serán

2, 8, 20, 40, 70, 112, ...

**Degeneración del espectro de energías**

$$g_n = (n + 1)(n + 2)$$

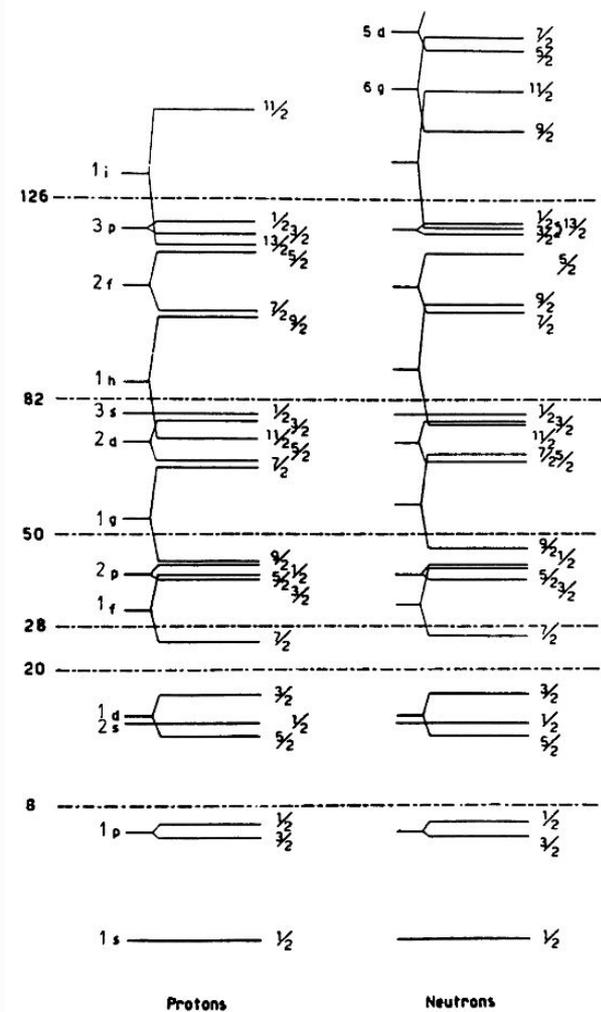
2, 6, 12, 20, 30, 42, ...

(en realidad esto va dividido por dos, pero luego se vuelve a multiplicar por dos por el spin de los fermiones)

# MODELO DE CAPAS

Un modelo más sofisticado para el campo medio es Woods Saxon + spin-orbit

$$U(r) = -\frac{V_0}{1 + e^{\frac{r-R}{a}}} + \xi \mathbf{s} \cdot \mathbf{l} \longrightarrow$$



# MODELO DE CAPAS

---

Se vuelve a la hamiltoniana many body

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{\mathcal{H}}_{res}$$

$$\hat{\mathcal{H}}\psi = E\psi$$

Donde  $\psi$  es la función de onda many body de los nucleones de valencia, que por ejemplo puede escribirse como un determinante de Slater de las SIP

# MODELO DE CAPAS

---

Proponiendo una energía residual 2 cuerpos, se puede llegar a

$$BE(j^n \text{ g.s.}) = BE(n = 0) + n\epsilon_j + \frac{n(n-1)}{2}\alpha + \frac{n - \frac{1}{2}(1 - (-1)^n)}{2}\beta$$

Por otros caminos se puede llegar a las energías de los estados excitados

Formalismo de número cuántico seniority ( $\nu$ )

# MODELO DE CAPAS

---

## A favor

- Es más fundamental, sólo se debe proponer un modelo para la interacción nucleón-nucleón, y las demás propiedades emergen como resultado
- Predice los números mágicos y los estados excitados

## En contra

- La predicción cuantitativa no se verifica experimentalmente para algunos núcleos lejos de capas cerradas
- El cálculo es complicado

02

---

# GENERALIDADES DE REDES NEURONALES

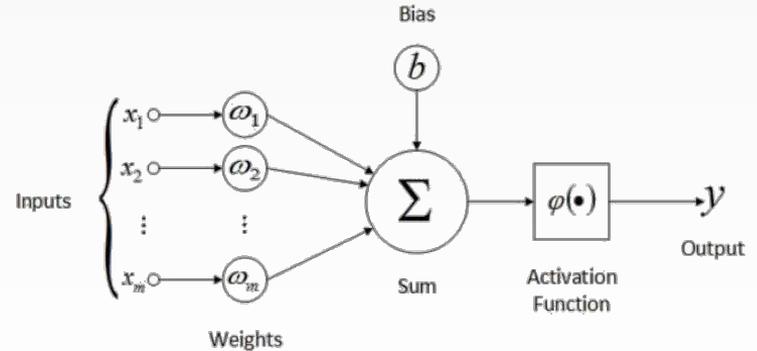
# ¿QUÉ SON LAS REDES NEURONALES?

Las redes neuronales se pueden entender como un tipo de funciones multivariable.

El valor de la función se calcula como una combinación de funciones lineales, y posiblemente aplicación de funciones no lineales sencillas sobre el resultado.

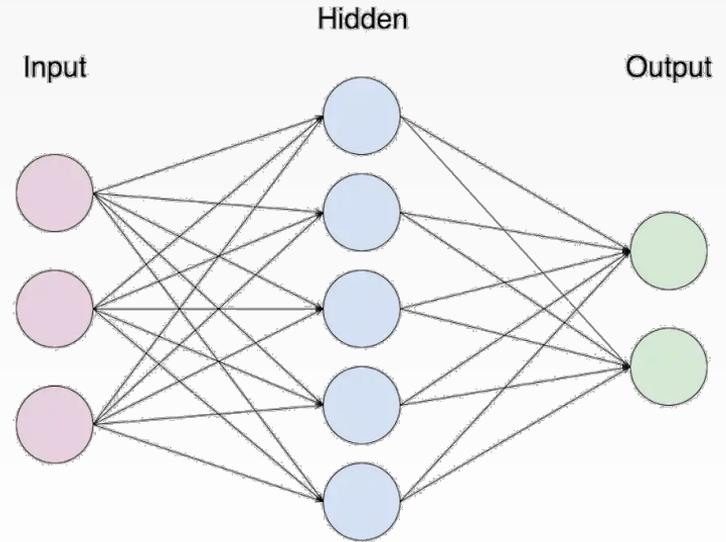
$$\Sigma_i = \sum_j \omega_{ij} x_j + b_i$$
$$y_i = \phi(\Sigma_i)$$

## Perceptrón

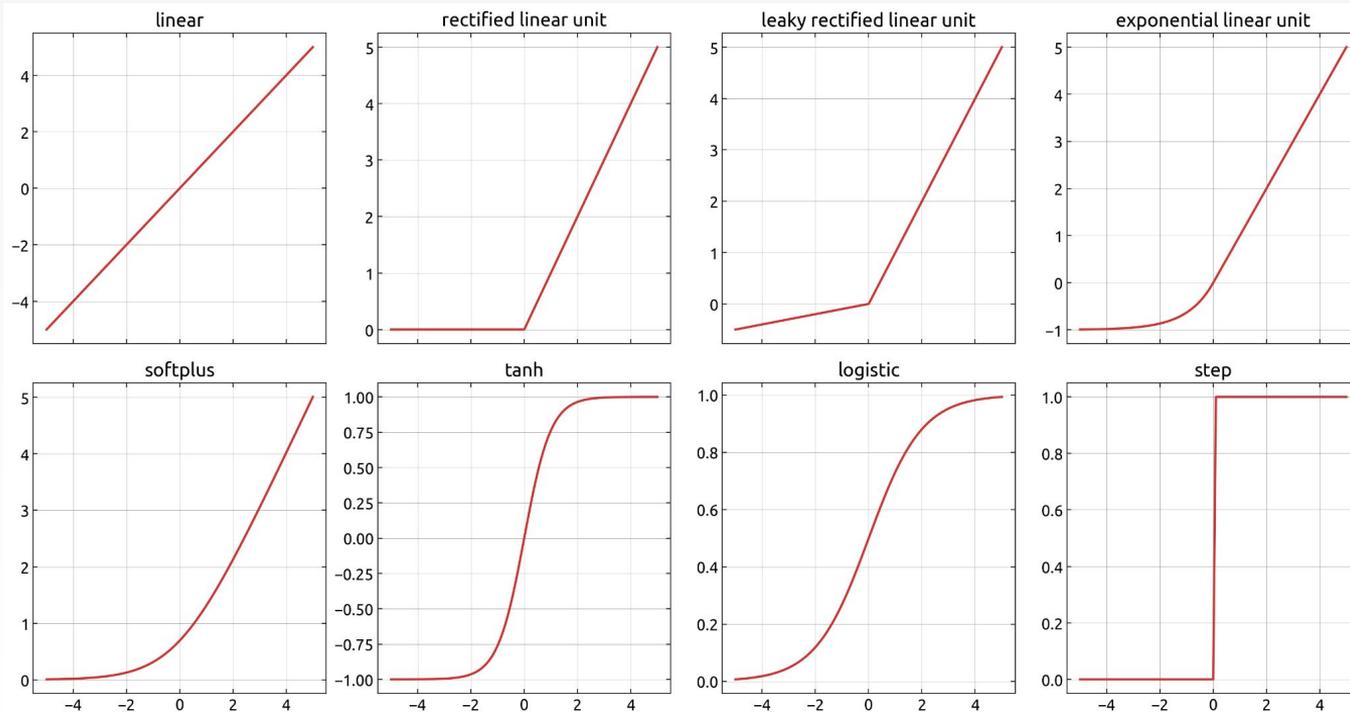


# REDES FEED FORWARD

Los perceptrones individuales se pueden acoplar en redes neuronales feed forward o perceptrones de muchas capas (multi-layer perceptron) para devolver respuestas más complejas frente a la entrada.



# FUNCIONES DE ACTIVACIÓN



# ¿POR QUÉ REDES NEURONALES?

---

Teorema de Kolmogorov, versión de Cybenko (1989)

- Cualquier función multivariable continua y real definida en un dominio compacto puede ser **aproximada con precisión arbitraria** por una red neuronal de una capa oculta con pesos adecuados y funciones de activación **sigmoideas** (ej: tanh, logística, etc).

# ENTRENAMIENTO O APRENDIZAJE

---

Para encontrar pesos adecuados proponemos una función error o 'loss' (pérdida), que será función de los mismos, y la optimizamos, proceso que denominamos 'entrenamiento'.

Un ejemplo básico es el problema de interpolación de datos, donde podemos usar la función de error de mínimos cuadrados

$$\mathcal{L} = \frac{1}{N_k} \sum_k (Y_k - f(X_k))^2$$

# ENTRENAMIENTO O APRENDIZAJE

---

Para optimizar la función de pérdida tenemos distintas opciones

- Usar un método de optimización multivariable general, como puede ser L-BFGS
- Usar métodos típicos de machine learning como pueden ser los métodos de descenso de gradiente estocástico

03

---

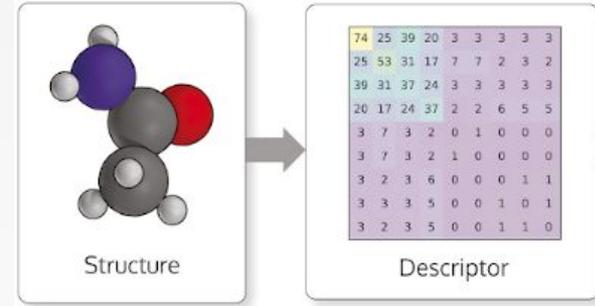
# COMPLEMENTOS DE RN

# FEATURE ENGINEERING

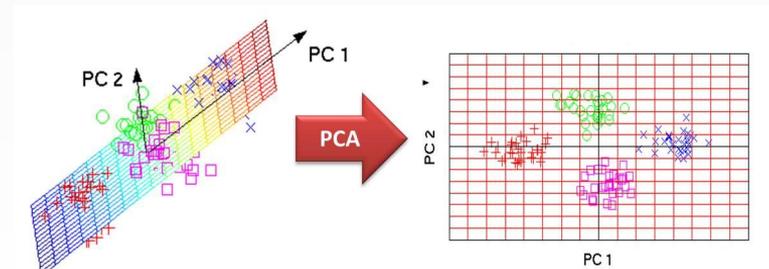
Los modelos de ML necesitan formatear su entrada, no funcionan bien usando los datos en crudo.

Aplicando conocimiento del área y/o métodos estadísticos se seleccionan transformaciones que lleven los datos en crudo a entradas útiles para el modelo.

## Descriptores del entorno químico



## Descomposición en ejes principales





# RESTRICCIONES Y REGULARIZACIÓN

---

El objetivo es que al entrenar modelos, se tengan en cuenta leyes físicas, propiedades matemáticas, u otros patrones que gobiernan el problema, forzando que nuestras posibles soluciones respeten un cierto comportamiento deseado.

La forma de hacerlo es agregar términos a la función loss que penalicen soluciones que no satisfagan las reglas que queremos imponer.



# EJEMPLO: MEJORAR LA SUAVIDAD DE LA FUNCIÓN

Un caso típico es incorporar la condición de continuidad de Lipschitz a un modelo de interpolación.

Se escribe una función loss total

$$\mathcal{L} = \mathcal{L}_{MSE} + \lambda \mathcal{L}_{Lips}$$

Donde  $\mathcal{L}_{MSE}$  es el error cuadrático medio usual,  $\lambda$  un peso elegido ad hoc, y

$$\mathcal{L}_{Lips} = \prod_n k_n$$

$$k_n = \max_{\mathbf{a} \neq 0} \frac{\|\mathbf{W}^n \mathbf{a}\|}{\|\mathbf{a}\|}$$

## Condición de Lipschitz

Existe un  $k$  tal que para todo par de  $x$

$$|f(x_1) - f(x_2)| < k|x_1 - x_2|$$

## Notas

Si  $f(x) = g(h(x))$

la constante de la composición es el producto de las individuales

Las funciones de activación tienen constantes conocidas

# FEATURE IMPORTANCE

La mejor interpretación de un modelo sencillo, como puede ser un modelo lineal, es el modelo en sí. Para modelos complejos, como una red neuronal, necesitamos de métodos especializados para interpretarlos.

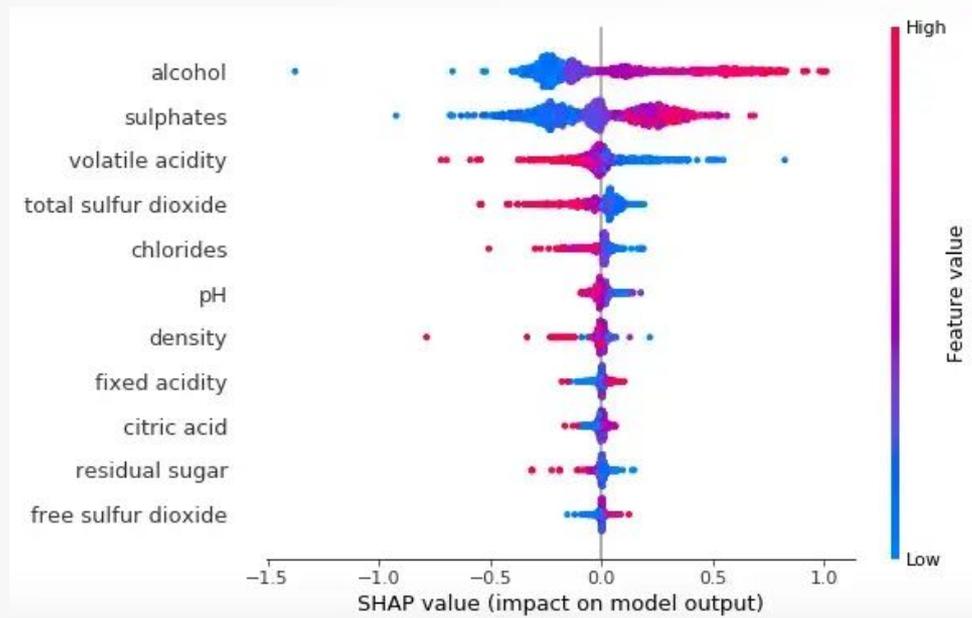
---

---

Feature importance son métodos de interpretación de modelos basados en asignar un índice de 'importancia' a cada parámetro de entrada en función del impacto que genera en cada predicción del modelo.

- Son fácilmente interpretables por las personas
- Dan una solución única (sería problemático que para un mismo modelo tengamos más de una posible interpretación)

## Ejemplo: SHAP VALUES en caracterización de calidad de vino



04

---

# PAPER DE MUMPOWER

- Como alternativa a los modelos semi empíricos y teóricos, los autores proponen un modelo de redes neuronales.
- El esquema que proponen es entrenar una red neuronal a 450 núcleos con  $Z \geq 20$ , con datos tomados de la base de datos AME 2016. (El set de entrenamiento es un ~20% de la base de datos)
- Emplean una red con arquitectura 6x8 y una capa de mezcla Gaussiana (devuelve una incerteza además del valor medio), y usan el optimizador de Adam (grad stoc)

## Features aumentadas

- Z
- N



- Z
- N
- $A = Z + N$
- $V_n$  : neutrones de valencia
- $V_z$  : protones de valencia
- $N_{eo}$  : 0 si N es par, 1 si N es impar
- $Z_{eo}$  : 0 si Z es par, 1 si Z es impar
- $P_{asym} = (N - Z)/A$

- La función loss es similar a mínimos cuadrados, pero normalizada por la incerteza de los datos. (Derivado de las funciones de likelihood gaussianas en estadística)
- Incorpora un término de restricción por información física (L1) y de regularización por decaimiento de pesos (L2)

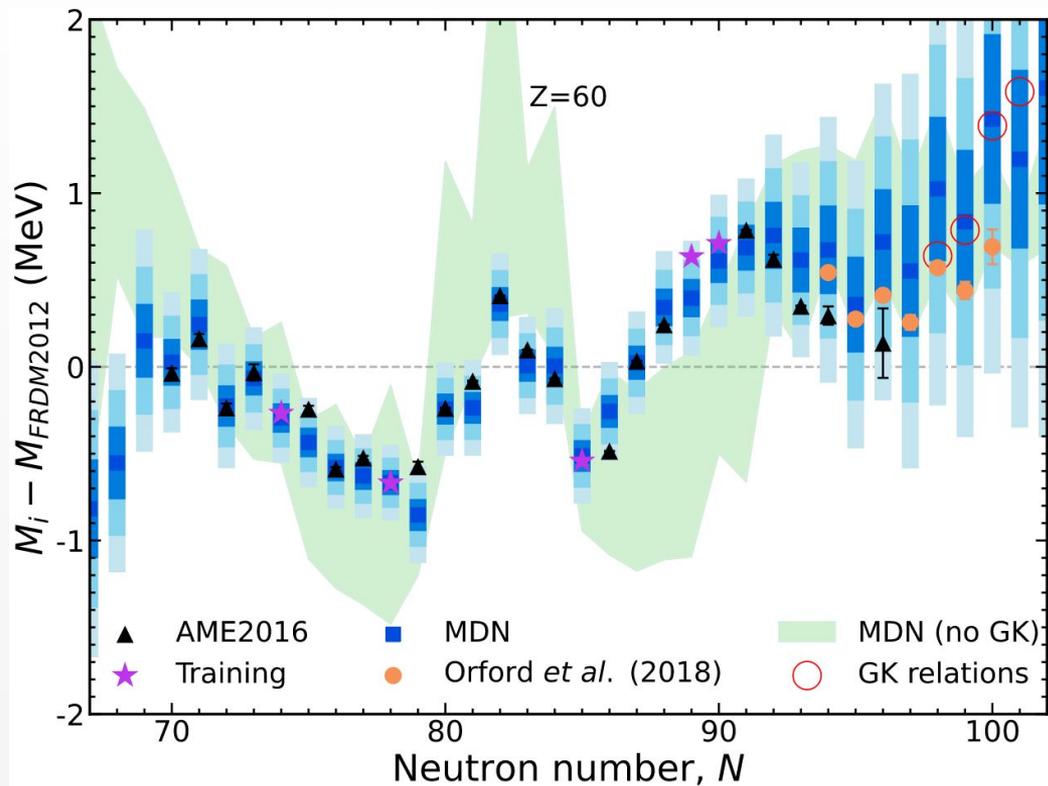
## Información física: relaciones de Garvey-Kelson

$$\begin{aligned} \mathbf{Z \leq N} \quad & M(Z - 2, N + 2) - M(Z, N) + M(Z - 1, N) \\ & - M(Z - 2, N + 1) + M(Z, N + 1) \\ & - M(Z - 1, N + 2) = 0 \end{aligned}$$

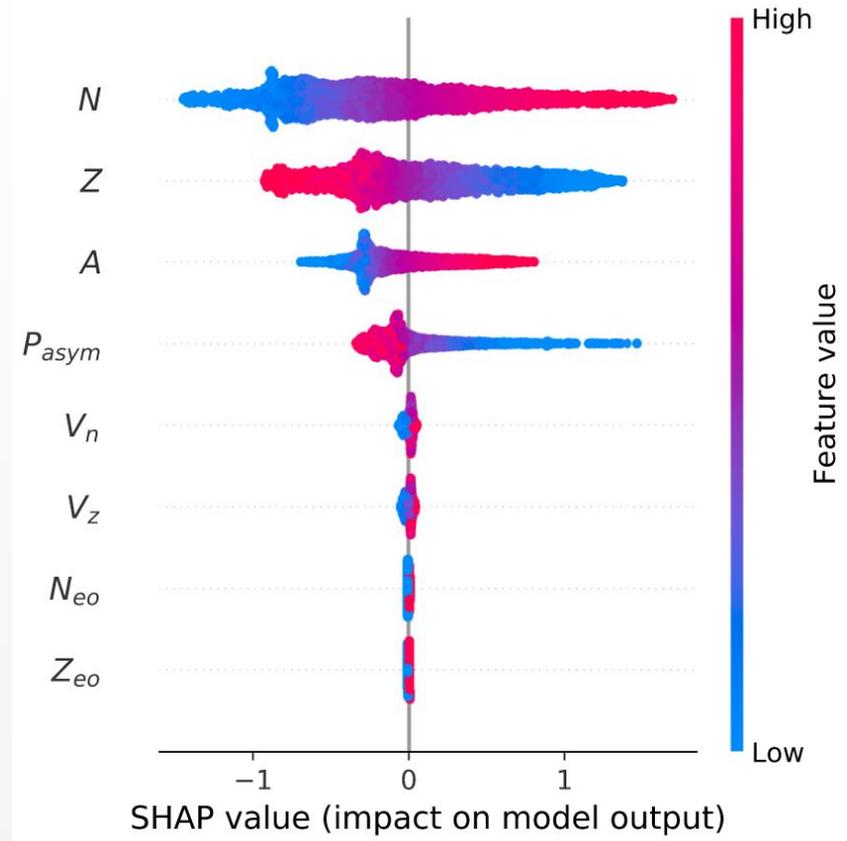
$$\begin{aligned} \mathbf{Z > N} \quad & M(Z + 2, N - 2) - M(Z, N) + M(Z, N - 1) \\ & - M(Z + 1, N - 2) + M(Z + 1, N) \\ & - M(Z + 2, N - 1) = 0 \end{aligned}$$

# Análisis de la cadena del Neodimio (Z=60)

- ▲ Base de datos (AME2016)
- ★ Muestras de entrenamiento
- Resultados de este trabajo con sus incertezas de 1, 2 y 3 sigma
- Resultados de trabajos previos de estos autores sin las GK
- Mediciones posteriores a la publicación de la base de datos
- Estimación usando las relaciones GK



- A primer orden dominan los parámetros “macroscópicos”  $N$ ,  $Z$  y  $A$ , seguidos por el parámetro de asimetría
- Se interpreta correctamente que los núcleos con  $Z > N$  tienden a ser más inestables
- El impacto de los parámetros “microscópicos” es menor, pero no nulo



# CONCLUSIONES

- El modelo da predicciones consistentes con la literatura
- Tiene utilidad para proveer extrapolaciones razonables en núcleos exóticos de vida corta, más aún provee incertezas confiables sobre la predicciones
- La metodología puede extenderse fácilmente para tener en cuenta otras propiedades de los núcleos

# REFERENCIAS

- Suhonen, J., From nucleons to nucleus. New York: Springer-Verlag Berlin Heidelberg. (2007)
- Heyde, K. L. G., The Nuclear Shell Model, Springer. (1994)
- Talmi, I. Simple models of complex nuclei. Harwood Academic Publishers. (1993)
- M. H. SAZLI, A brief review of feed-forward neural networks, Communications, Faculty Of Science, University of Ankara 11 (2006).
- V. K. Ojha, A. Abraham, and V. Snášel, Metaheuristic Design of Feedforward Neural Networks: A Review of Two Decades of Research, Engineering Applications of Artificial Intelligence 60, 97 (2017).
- Hassoun, Mohamad H., Fundamentals of Artificial Neural Networks, MIT Press, (2003).
- G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Physics-Informed Machine Learning, Nat Rev Phys 3, 422 (2021).
- M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations, Journal of Computational Physics 378, 686 (2019).

# REFERENCIAS

- S. M. Lundberg, and S. Lee, A unified approach to interpreting model predictions, Proceedings of the 31st International Conference on Neural Information Processing Systems, 4768–4777 (2017).
- J. Mendoza-Temis, A. Frank, J. G. Hirsch, J. C. López Vieyra, I. Morales, J. Barea, P. Van Isacker, and V. Velázquez, Nuclear Masses and the Number of Valence Nucleons, Nuclear Physics A 799, 84 (2008).
- M. R. Mumpower, T. M. Sprouse, A. E. Lovell, and A. T. Mohan, Physically Interpretable Machine Learning for Nuclear Masses, Phys. Rev. C 106, (2022).